Proposal: A Twist on Adaptive Metropolis-Hasting Algorithm

Ivy Liu, Zefan Liu

2024-03-20

1. Introduction

Markov Chain Monte Carlo (MCMC) methods are a class of algorithms for sampling from probability distributions based on constructing a Markov chain that has the desired distribution as its equilibrium distribution. As the chain progresses, MCMC provides a way to estimate the posterior distribution of the parameters.

Traditional MCMC methods include the Metropolis-Hastings algorithm and the Gibbs sampler. The effectiveness (the convergence rate) of traditional MCMC methods heavily depends on the choice of the proposal distribution and tuning parameters, such as the step size. These parameters need to be carefully selected and fixed throughout the sampling process. If poorly chosen, the chain might require a long time to converge to the target distribution, leading to inefficient sampling.

Alternatively, adaptive MCMC methods aim to overcome the limitations of traditional MCMC by allowing the algorithm to automatically adjust its parameters during the sampling process (Andrieu & Thoms, 2008). Specifically, the AM algorithm (Haario et al., 2001) aims to use a proposal distribution that is adjusted by the empirical estimate of the covariance structure of the target distribution based on the iterations so far. In this project, we will propose a twist on the AM algirithm, such that its step size is determined also by the acceptance ratio of the previous run.

2. Methodology

2.1 Algorithm

We propose the following Markov Chain Monte Carlo algorithm, which imposes a twist on the proposal structure of Adaptive Monte Carlo (AMC) method (Andrieu & Thoms, 2008). The overall idea is still updating the proposal at each step of sampling based on a combination of empirical posterior distribution and a baseline proposal distribution, where the weight from the baseline proposal is relatively small, but our twist will also update the baseline proposal based on the previous stepsize and likelihood ratio. Note that some initialized values are following the convention in previous works (Andrieu & Thoms, 2008).

2.2 Proof of Ergodicity

To prove that our Adaptive MCMC method will vanish, we need to prove our algorithm satisfy the following two conditions (Roberts & Rosenthal, 2007)

1. **Diminishing Adaptation**: $\lim_{n \to \infty} \sup_{x \in \mathbf{X}} ||P_{\Gamma_{n+1}}(x) - P_{\Gamma_n}(x)||$ is $o_p(1)$

2. **Bounded Convergence**: The time that distribution that our algorithm preserves converges to the target distribution is finite.

Algorithm 1 Let's try

1: Initialize $\Gamma_0 = \mathbf{I}_d, \beta = 0.05, \alpha_0 = 0.4$, and $\sum = \mathbf{I}_d$. Let N denotes the iteration times, x denotes samples and $\gamma(x)$ denotes the unnormalized density of the distribution to sample from. Let x_0 denote the initial value picked from the target distribution.

2: For $n = 1, 2, ..., N$, repeat

3: **if** $i \geq 2$ **then**

4:     $\Sigma \leftarrow \frac{1}{n} \sum_{i=1}^{n-1} (x_i - \bar{x})(x_i - \bar{x})^T$ ▷ covariance matrix

5: **end if**

6: $W_n \leftarrow \frac{\alpha_{n-1} - 0.4}{\sqrt{n}}$ ▷ Step size adjustment

7: $\Gamma_n \leftarrow (1 + W_n)\Gamma_{n-1}$

8: $x' \sim Q_n(x_n|x_{n-1}) = (1 - \beta)N(x_{n-1}, (2.38)^2\Sigma_n/d) + \beta N(x_{n-1}, (0.1)^2\Gamma_n/d)$

9:   ▷ Propose x' from the mixture proposal

10: $\alpha_n \leftarrow min(1, \frac{\gamma(x')Q_n(x_{n-1}|x')}{\gamma(x_{n-1})Q_n(x'|x_{n-1})})$ ▷ The Hastings ratio

11: $u \sim Unif(0,1)$

12: **if** $u \leq \alpha_n$ **then**

13:     $x_n \leftarrow x'$

14: **else**

15:     $x_n \leftarrow x_{n-1}$

16: **end if**

17: Return a sequence of samples $\{x_i\}_{i=1}^N$.

2.3 Simulation

In this section, we will evaluate the performance of our proposed algorithm, the AM algorithm, and the traditional MH algorithm on the following distribution:

1. The standard normal distribution

2. The mixture of two normal distribution

3. A funnel-shaped 2D distribution

4. A banana-shaped 2D distribution

References

Andrieu, C., & Thoms, J. (2008). A tutorial on adaptive MCMC. *Statistics and Computing, 18, issue 4*, 343–373. https://doi.org/10.1007/s11222-008-9110-y

Haario, H., Saksman, E., & Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli, 7*(2), 223–242.

Roberts, G. O., & Rosenthal, J. S. (2007). Coupling and ergodicity of adaptive markov chain monte carlo algorithms. *Journal of Applied Probability, 44*(2), 458–475. http://www.jstor.org/stable/27595854